MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo 357 May 1976

FROM UNDERSTANDING COMPUTATION TO UNDERSTANDING NEURAL CIRCUITRY

by

D. Marr and T. Poggio*

Abstract: The CNS needs to be understood at four nearly independent levels of description: (1) that at which the nature of a computation is expressed; (2) that at which the algorithms that implement a computation are characterized; (3) that at which an algorithm is committed to particular mechanisms; and (4) that at which the mechanisms are realized in hardware. In general, the nature of a computation is determined by the problem to be solved, the mechanisms that are used depend upon the available hardware, and the particular algorithms chosen depend on the problem and on the available mechanisms. Examples are given of theories at each level.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-75-C-0643.

* Permanent address: Max-Planck Institut fur Biologische Kybernetik,
74 Tubingen 1, Spemannstrasse 38, Germany.

From understanding computation to understanding neural circuitry

Complex systems, like a nervous system or a developing embryo, must be analyzed and understood at several different levels. Of course, there are logical and causal relationships among them, but the important point is that these levels of description are only loosely related. The underlying philosophical issue here is that reductionism does not imply constructionism.

Thus the question we ask in this essay is, at what level is it necessary and profitable to study the information processing that is carried out during visual perception? For a machine that solves an information processing problem, there are four important levels of description. At the lowest, there is basic component and circuit analysis -- how do transistors, neurons, diodes, and synapses work? The second level is the study of particular mechanisms; adders, multipliers and memories accessed by address or by content. The third level is that of the algorithm, and the top level contains the theory of the overall computation. For example, take the case of Fourier analysis. The theory of the Fourier transform is well understood, and is expressed independently of the particular way in which it is computed. One level down, there are several algorithms for implementing a Fourier transform -- the Fast Fourier transform (Cooley & Tukey 1965) which is a serial algorithm; and the parallel algorithms of holography that are based on mechanisms of laser optics. All these algorithms carry out the same computation, and the choice of which one to use depends upon the particular mechanisms that are available. If one has fast digital memory, adders and multipliers, one will use the FFT, and if one has a laser and photographic plates, one will use an optical algorithm. In general, mechanisms are strongly determined by hardware, the nature of the computation is determined by the problem, and the algorithms are determined by the computation and the available mechanisms.

Each level of description has its place in the eventual understanding of perceptual information processing, and it is important to keep them separate. Too often in attempts to relate psychophysical problems to physiology there is confusion about the level at which a problem arises -- is it related mainly to biophysics (like after-images) or primarily to information processing (like the ambiguity of the Necker cube)? More disturbingly, although the top level is the most neglected, it is also the most important. This is because the structure of the computations that underly perception depend more upon the computational *problems* that have to be solved than on the particular hardware in which their solutions are implemented. There is an analog of this in physics, where the global descriptions of thermodynamics represented, at least historically, the first stage in the study of matter. A description in terms of mechanisms or elementary components appeared afterwards.

Our main point then is that the topmost of our four levels, that at which the necessary structure of computation is defined, is a crucial but neglected one. Its study is separate from the study of particular algorithms, mechanisms or hardware, and the techniques needed to pursue it are new. In the rest of the article, we summarize some examples of theories at the different levels we described, and illustrate the types of

Figure 1. 1a shows the distributions of the error angle $\psi(t)$, during stationary fixation of 2-stripe patterns (lower figure). The varying parameter is the angular separation of the two black, vertical stripes. The upper figure shows corresponding histograms obtained from eq. (1). A typical phase transition occurs in the stationary fixation distribution for a value of the parameter between 40 and 60 degrees. The nonlinear superposition of very simple, local mechanisms (see Fig. 5b) leads to such a symmetry breaking. An open loop analysis (for instance *via* electophysiology) cannot predict this closed loop behavior without the phenomenological theory. (From Reichardt and Poggio, 1976). The same phase transition behavior can be observed in the fixation of the Mueller-Lyer figures (fig.1b). The histograms extend from -180 to 180 degrees and show the fraction of time the fly gazed at any part of the two figures. The results conform to the theory's predictions (from Geiger and Poggio, 1975). A fly fixates (in the horizontal degree of freedom) the "illusory" vertical line arising at the boundary between the two sets of parallel horizontal lines in the upper pattern (fig. 1c). There is no fixation, however, of a similar illusory, vertical line in the lower pattern. The phenomenological theory correctly predicts these results. (From Poggio and Geiger, unpublished).

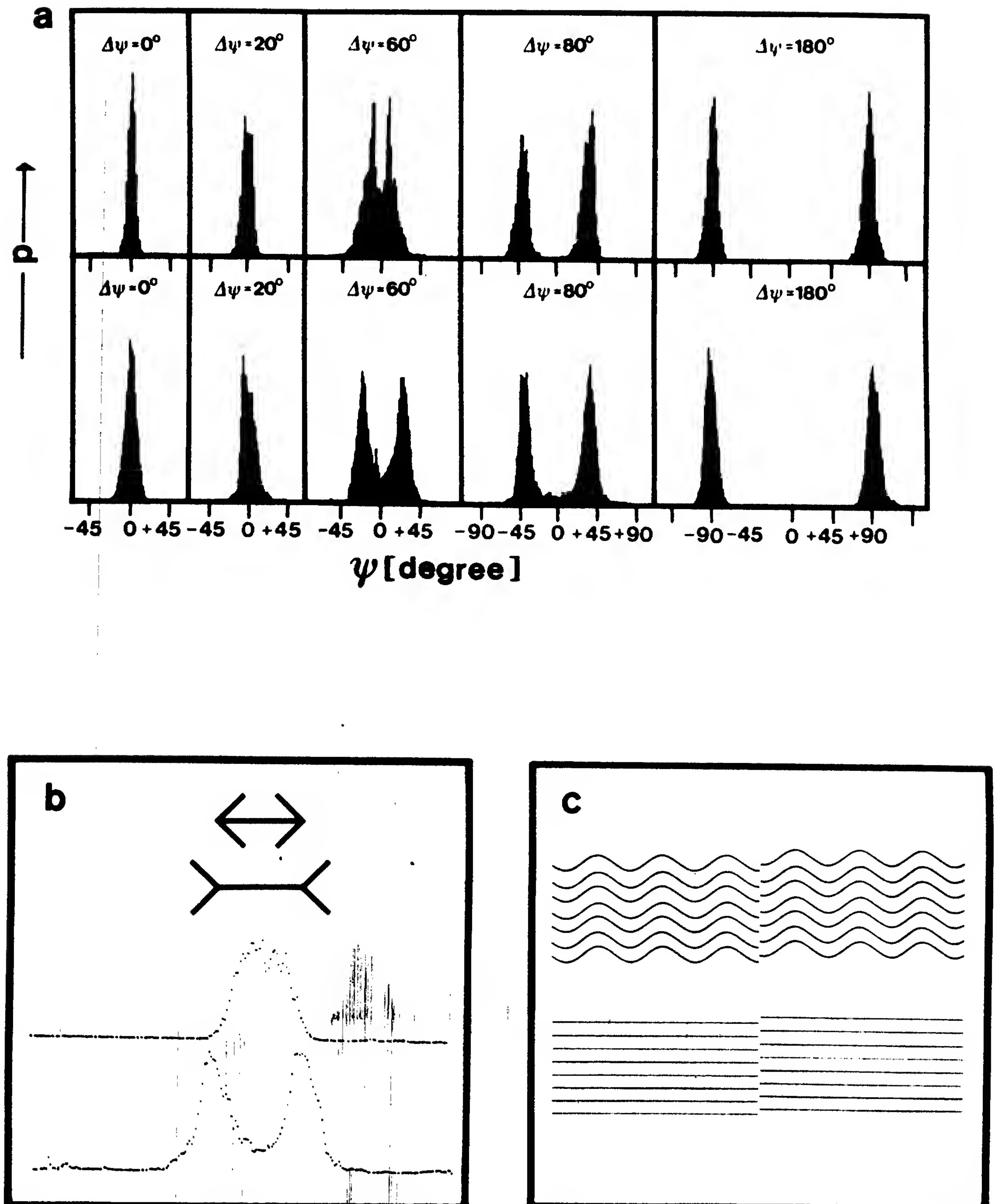
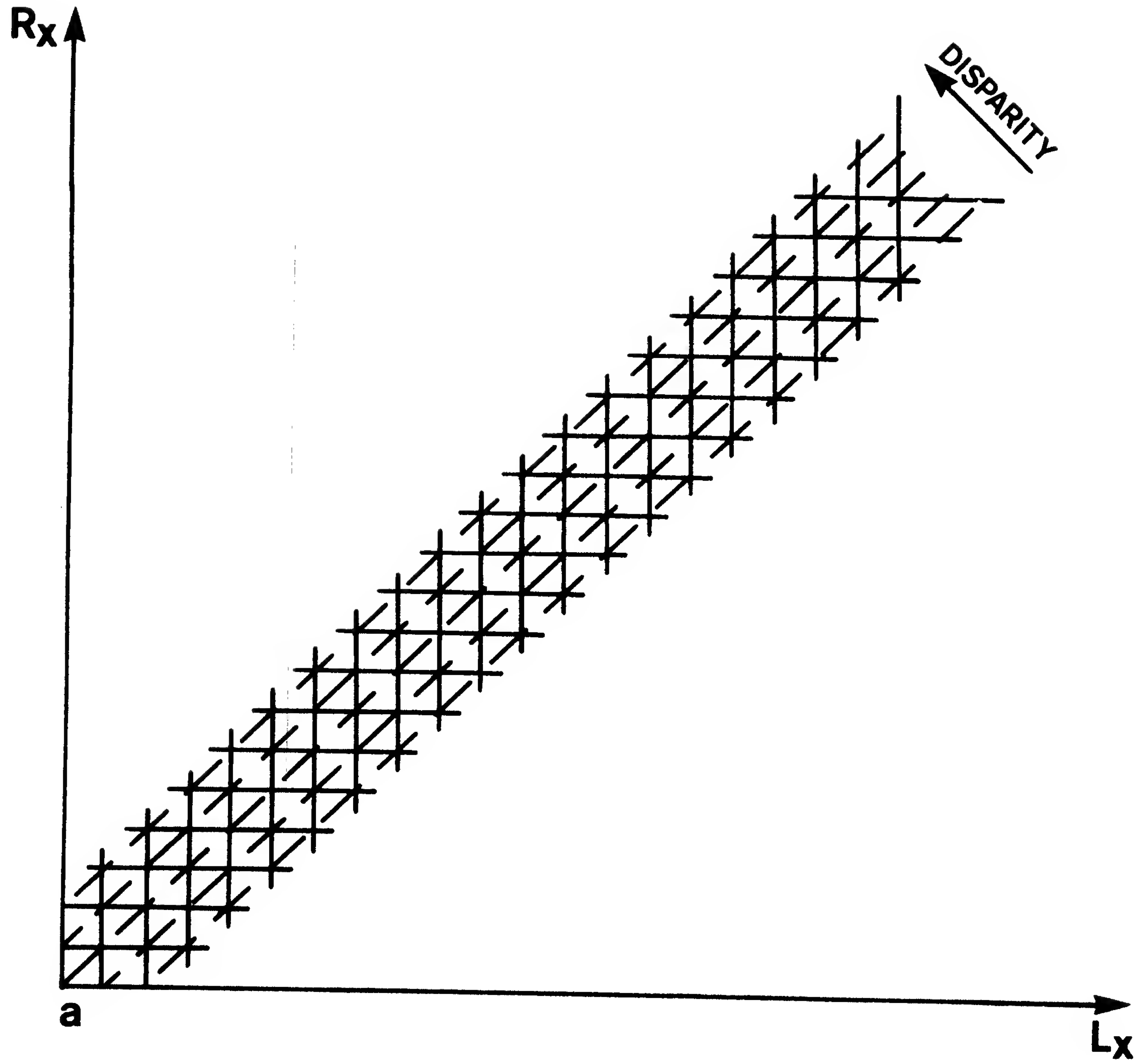

a
Δψ=0°
Δψ=20°
Δψ=60°
Δψ=80°
Δψ=180°
Δψ=0°
Δψ=20°
Δψ=60°
Δψ=80°
Δψ=180°
p
-45 0 +45 -45 0 +45 -45 0 +45 -90 -45 0 +45 +90 -90 -45 0 +45 +90
ψ [degree]
b
c

FIGURE 1

Figure 2. The geometry of constraints on the computation of binocular disparity. 2a illustrates the constraints for the case of a one-dimensional image. Lx and Ly represent the positions of descriptive elements from the left and right views, and the horizontal and vertical lines indicate the range of disparity values that can be assigned to left-eye and right-eye elements. The use-once condition states that only one disparity value may be assigned to each descriptive element. That is, only one disparity value may be "on" along each horizontal or vertical line. The second condition states that we seek solutions in which disparity values vary smoothly almost everywhere. That is, solutions tend to spread along the dotted diagonals, which are lines of constant disparity, and between adjacent diagonals. 2b shows how this geometry appears at each intersection point. The constraints may be implemented by a network with positive and negative interactions that obey this geometry, because the stable states of such a network are precisely the states that satisfy the constraints on the computation. 2c shows the constraint geometry for a 2-dimensional image. The negative interactions remain essentially unchanged, but the positive ones now extend over a small 2-dimensional neighbourhood. A network with this geometry was used to perform the computation exhibited in figure 6.



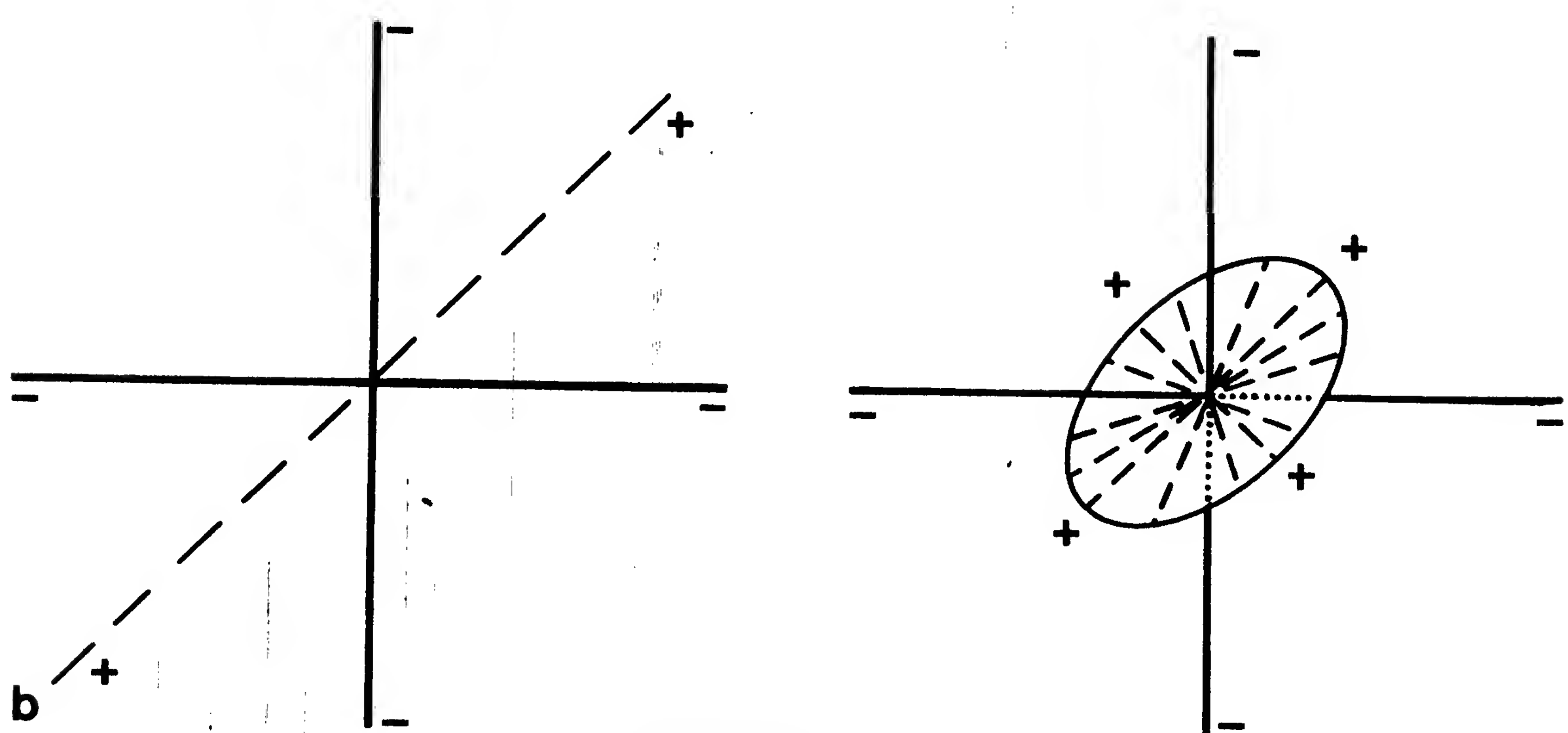


FIGURE 2

prediction that can emerge from each.

Examples of computational theories

1: Orientation behavior of the fly (Reichardt, 1970; Reichardt & Poggio, 1976)

The approach of Reichardt and Poggio towards an understanding of the visual control system of the fly is an example of the characterization of a "simple" complex system. Their description accounts in a quantitative way for orientation, chasing and spontaneous pattern preference behavior. Although connections with physiological and anatomical data are being established, the theory was based on behavioral data. The theory outlines the basic logical organization of the visual control system of the fly. It holds that the nervous system performs two main computations on the visual input, one extracting movement information, the other providing position information. The theory leads to the equation of motion

$$\Theta\ddot{\psi}(t) + k\dot{\psi}(t) + k\omega(t) = -D[\psi(t)] - r\dot{\psi}(t) + N(t) \qquad (1)$$

where the angular error $\psi(t)$ represents the instantaneous position of the pattern on the retina of the fly. The terms on the left hand side represent the flight dynamics (Θ is the moment of inertia of the fly, k is a rotational friction constant). $\omega(t)$ is the angular speed of the object. The right hand side describes the instantaneous torque of the fly; the term $N(t)$ is a zero-mean random process, and is independent of the visual input; $r\dot{\psi}(t)$, a velocity-dependent optomotor response, is the result of "movement computation"; $D(\psi)$ carries the position information, acquired from the visual input by the "position computation". All these terms have been characterized quantitatively, through independent experiments. Through equation (1) the theory predicts a rather complex natural (closed loop) behavior, characterized by phase transition-like phenomena, and by primitive classifications of patterns. Figure 1 gives two examples of behavior that is quantitatively explained by this approach. The quantitative description of equation (1) could not have been obtained from single cell recordings or from histology. Furthermore, it is probably a prerequisite of any full understanding at the level of circuitry.

The other examples that we describe come from work on visual information processing that has been carried out at the M. I. T. Artificial Intelligence Laboratory over the last two years.

2: Stereopsis (Marr 1974, Marr & Poggio in preparation)

Suppose that images of a scene are available, taken from two nearby points at the same horizontal level. In order to compute stereoscopic disparity, the following steps must be carried out: (1) a particular location on a surface in the scene must be chosen from one image; (2) that location must be identified in the other image; (3) the relative positions of the two images of that location must be measured. Notice that methods based on grey-level correlation between images fail to satisfy these conditions because a grey-level measurement does not define a point on a physical surface independently of the geometry of the imaging device. The matching must be based on objective markings that lie on a

physical surface, and so one has to use predicates that correspond to changes in reflectance. One way of doing this is to obtain a primitive description of the intensity changes that exist in each image, and then to match these descriptions. Line and edge segments, blobs, and edge termination points correspond quite closely to boundaries and reflectance changes on physical surfaces.

The stereo problem may thus be reduced to that of matching two primitive descriptions, one from each eye. One can think of elements of these descriptions as having only position information, like the black points in a random-dot stereogram, although in practise there exist some rules about which matches between descriptive elements are possible, and which are not. There are two crucial constraints on the way in which the left and right descriptions are combined:

(1) *The use-once condition*. Each item from each image may be assigned at most one disparity value. This condition rests on the premise that the items to be matched have a physical existence, and can be in only one place at a time.

(2) *Continuity*. Disparity varies smoothly almost everywhere. This condition is a consequence of the cohesiveness of matter, and it states that only a relatively small fraction of the area of an image is composed of boundaries.

These conditions on the computation are represented geometrically in figure 2a. Later in the article, we exhibit a network that implements these conditions, and we illustrate how it solves random-dot stereograms.

3: The Primal Sketch (Marr 1976a)

It is a commonplace that a scene and a drawing of the scene appear very similar, despite the completely different grey-level images to which they give rise. This suggests that the artist's local symbols correspond in some way to natural symbols that are computed out of the image during the normal course of its interpretation. The first part of this visual information theory asserts that the first operation on an image is to transform it into its raw *primal sketch*, which is a primitive but rich *description* of the intensity changes that are present. Figure 3 shows an example. In order to obtain this description, approximations to the first and second directional derivatives of intensity are measured at several orientations and on several scales everywhere in the image, and these measurements are combined to form local descriptive assertions. The process of computing the primal sketch involves five important steps, the first of which can be compared with the measurements that are apparently made by simple cells in the visual cortex. One prediction made by this part of the theory is that a given intensity change itself determines which simple-cell measurements are used to describe it. This is in direct contrast to theories which assert that each simple cell acts as a "feature-detector", whose output is freely available to subsequent processes. Another prediction is that a well-defined interaction must take place between simple-cell like measurements made at the same orientation and position in the visual field but with different receptive field sizes.

4: Grouping and texture vision (Marr 1976a)

The primal sketch of an image is in general a large and unwieldy

Figure 3. 3a shows the image of a toy bear, printed in a font with 16 grey levels. In 3b, the intensity at each point is represented along the z-axis. 3c illustrates the spatial component of the raw primal sketch as obtained from this image. Associated with each line segment are measures of contrast, type and extent of the intensity change, position and orientation. This image is so simple that purely local grouping processes suffice to extract the major forms from the primal sketch. These forms are exhibited in 3d, e & f. (from Marr 1976a, figure 21).

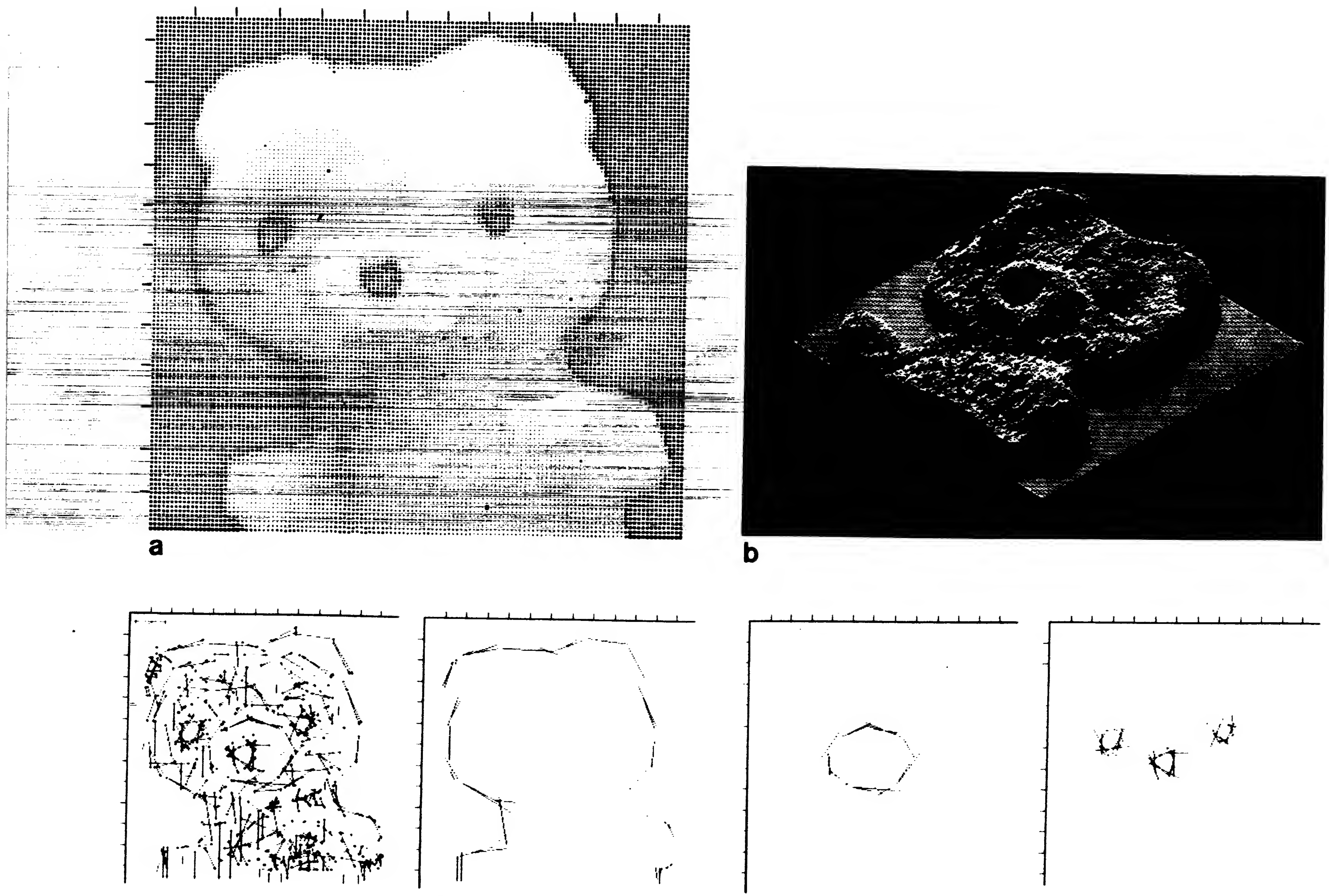


FIGURE 3

collection of data. This is an unavoidable consequence of the irregularity and complexity of natural images. The next important computational problem is how to decode the primal sketch. For most images, it is unnecessary to invoke specific hypotheses about what is there until considerably later in the processing. The theory next applies a number of quite general selection and grouping processes to elements in the primal sketch. The purpose of these processes is to organize the local descriptive elements into *forms* and *regions*, which are closed contour groups that are obtained in various ways. Regions may be defined by their boundaries, which have been formed by grouping together some set of edge, line, or place-tokens; or they may be defined by a first-order predicate operating on the primal sketch elements within it. This second method corresponds to the definition of a region by a texture, and it leads to a theory of the processes on which texture discrimination is based.

It is important to realize that the descriptive items that may be grouped here can be very abstract -- like tokens for the end of a line, a blob, or a constructed line that joins two blobs. Tokens are created for each new group, and these tokens themselves become subject to the operation of the same or similar grouping processes as operated on elements of the raw primal sketch. The grouping processes are very conservative. They satisfy a principle that seems to have general application to recognition problems, called the *principle of least commitment*. Only "obvious" groupings are made, and where there is doubt between two possible groupings, both are constructed and held pending subsequent selection. Figure 3 illustrates some results of applying these grouping processes.

5: 3-D representation of shape (Marr & Nishihara 1975)

The last two components of the theory concern the representation of three-dimensional shapes. One component deals with the nature of the representation system that is used, and the other with how to obtain it from the types of description that can be delivered from the primal sketch. The key ingredients of the representation system are:

(a) The deep structure of the three-dimensional representation of an object consists of a stick figure, where in formal terms each stick represents one or more axes in the object's generalized cylinder representation. In fact, a hierarchy of stick figures exists, that allows one to describe an object on various scales with varying degrees of detail.

(b) Each stick figure is defined by a propositional database called a *3-D model*. The geometrical structure of a 3-D model is specified by storing the relative orientations of pairs of connecting axes. This specification is local rather than global, and it contrasts with schemes in which the position of each axis is specified in isolation, using some circumscribing frame of reference.

(c) When a 3-D model is being used to interpret an image, the geometrical relationships in the model are interpreted by an essentially analogue mechanism called the *image-space processor*. This mechanism is computationally very simple, and may be thought of as a limited resolution device for representing the positions of two vectors in 3-space, and for computing their projections onto the image.

(d) During recognition, a sophisticated interaction takes place between the image, the 3-D model, and the image-space processor. This interaction gradually relaxes the stored 3-D

model onto the axes computed from the image. Some facets of this process resemble the computation of a 3-D rotation, but a simple computer graphics metaphor is misleading. In fact, the rotations take place on abstract vectors (the axes) that are not even present in the original image; and at any moment, only two such vectors are explicitly represented.

The essence of this part of the theory is a method for representing the spatial disposition of the parts of an object and their relation to the viewer.

6: 2 1/2 - dimensional analysis of an image (Marr & Vatan, in preparation)

In simple images, the forms delivered from the primal sketch correspond to the contours of physical objects. Finally therefore, we need to bridge the gap between such forms and the beginning of the 3-D analysis described in the previous paragraph. We call this 2 1/2 - dimensional analysis, and it consists largely of assigning to contours labels, that reflect aspects of their 3-dimensional configuration, before that configuration has been made explicit. The most powerful single idea here is the distinction between convex and concave edges and contour segments. One can show that these distinctions are preserved by orthogonal projections, and can be made the basis of a segmenting technique that decomposes a figure into 2-D regions that correspond to the appropriate 3-D decomposition for a wide range of viewing angles (see figure 4). The theory assigns many alternating figure effects like the Necker cube to the existence of alternative self-consistent labellings computed at this stage.

It is perhaps worth mentioning one interesting point that has emerged from this way of recognizing and representing 3-D shapes. Warrington & Taylor (1973) described patients with right parietal lesions who had difficulty in recognizing objects seen in "unconventional" views - like the view of a water pail seen from above. They did not attempt to define which views are unconventional. According to our theory, the most troublesome views of an object will be those in which its stick-figure axes cannot easily be recovered from the image. The theory therefore predicts that unconventional views in the Warrington & Taylor sense will correspond to those views in which an important axis in the object's generalized cylinder representation is foreshortened. Such views are by no means uncommon - if a 35mm camera is directed towards you, you are seeing an unconventional view of it, since the axis of its lens is foreshortened. Recent observations by S. Carey (personal communication) appear to confirm that this definition captures the important distinction.

Examples of algorithms and mechanisms

Between the top and bottom of our four levels lie descriptions of algorithms and descriptions of mechanisms. The distinction between these two levels is rather subtle, since they are often closely related. The form of a specific algorithm can impose strong constraints on the mechanisms, and conversely. Let us consider three examples.

Figure 4. Analysis of a contour. The outline (4a) was obtained from a primal sketch just as figure 3d was obtained from 3a. This contour is smoothed, and then divided into convex and concave components (3b). The outline is searched for deeply concave points or segments, which correspond to main segmentation points. There are usually several possible matches for each such segmentation point, but the correct mates for each may be found by eliminating relatively poor candidates. The result of this is the segmentation shown in 3c. Once these segments have been defined, corresponding axes are easy to obtain (3d). They do not usually connect, but may be related to one another by intermediate lines which are called embedding relations (3e). The resulting stick figure is shown in 3f which, according to the theory, is the deep structure on which interpretation of this image is based.

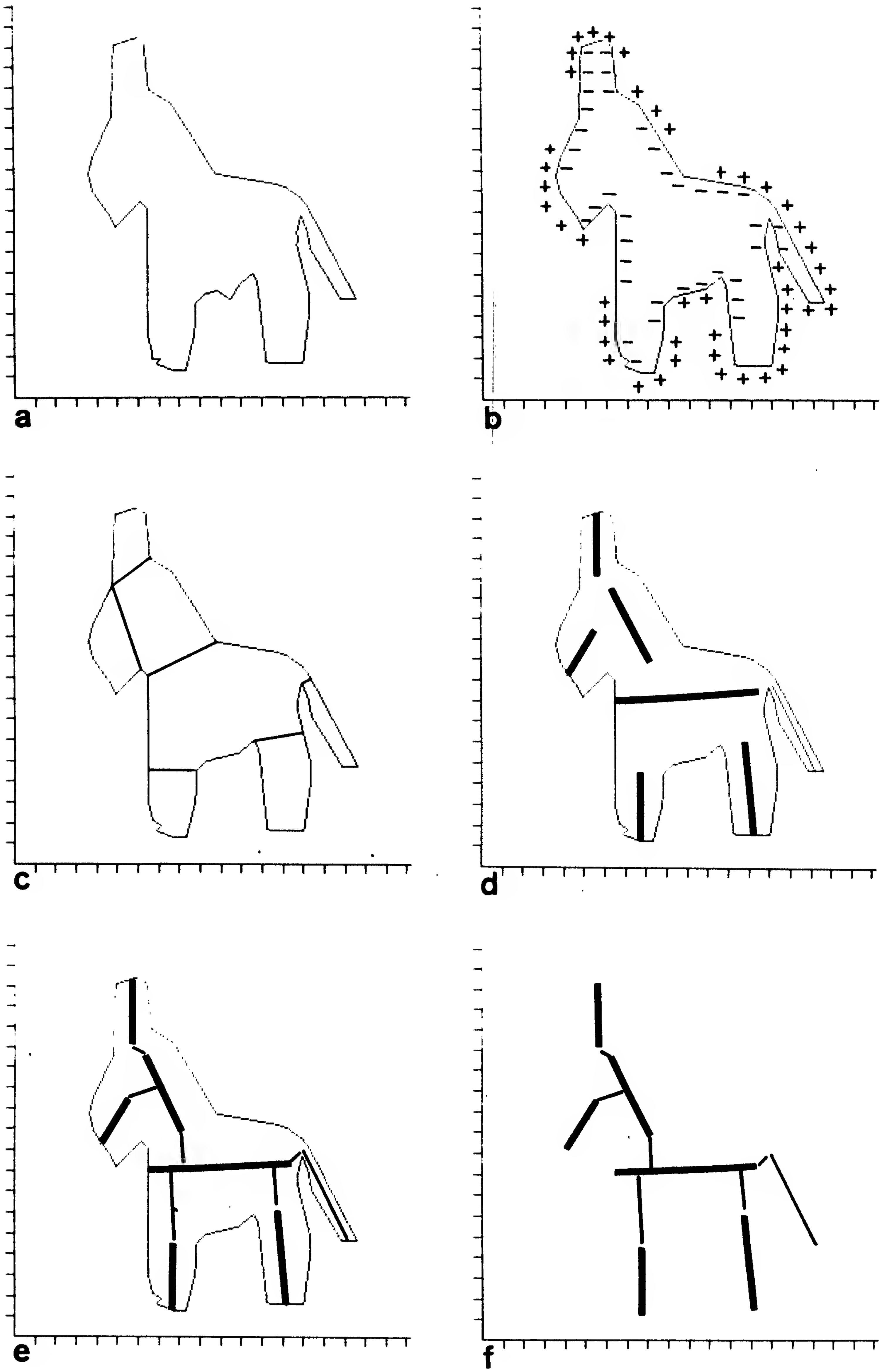


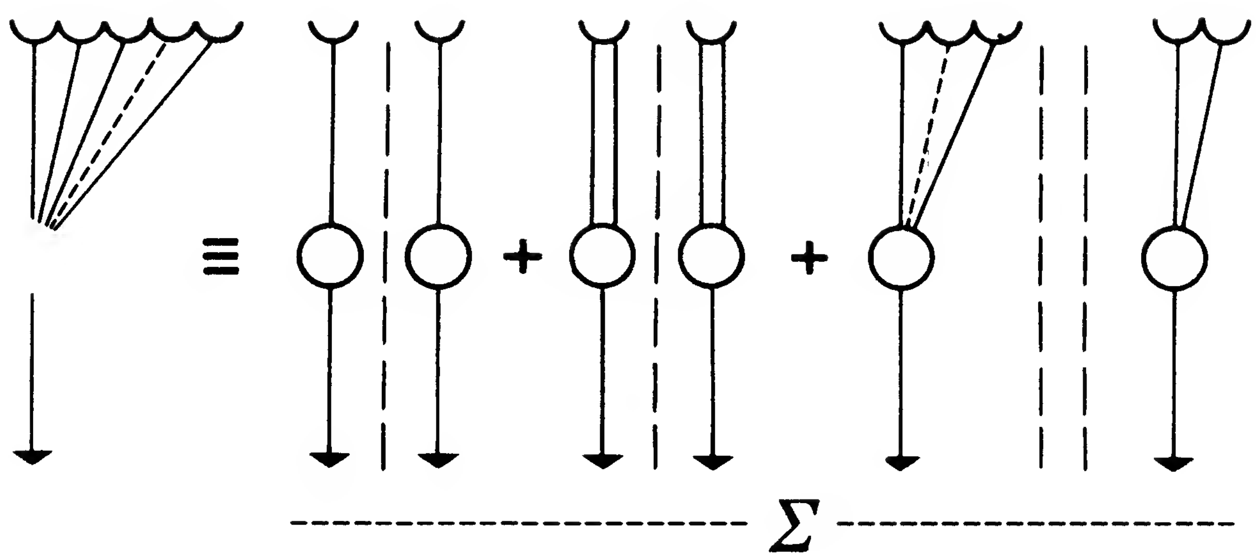
FIGURE 4

Figure 5. Graphical representation (a) of the decomposition of a nonlinear, n-input "algorithm" into a sum of interactions of various order. The functional representation

$$S\{\,..\,x_i(t)\,..\,\} = L^{(0)} + \sum L_i^{(1)}\,\{x_i(t)\} + \sum L_{ij}^{(2)}\,\{x_i(t), x_j(t)\} + \ldots$$

where $L^{(n)}$ is an n-linear mapping, can be read from an appropriate sequence of such elementary graphs. Fig. 5b shows the graphs that implement the fly's orientation behavior, studied by Reichardt and Poggio. Several findings suggest that they may correspond to separate physiological modules. Characteristic functional and computational properties can be associated to each interaction type. (From Poggio and Reichardt, 1976).

a



Separation of the three types of interactions in the fly

b

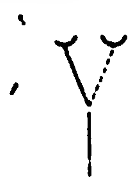
Movement computation	Position ("attractiveness") computation	
		
Corresponding to $r\dot{\psi}$	Corresponding to $D(\psi)$	Correction to superposition rule
Homogeneously distributed in the eye (no <u>strong</u> dependence on ψ and ϑ)	Mostly in the lower part of the eye ($D(\psi)$ and $L(\vartheta)$ dependence)	Mostly in the lower part of the eye
No "age" dependence	(?)	"Age" dependence
Light intensity threshold at about 10^{-4} candel/m^2 (Eckert, 1973)	Light intensity threshold (of fixation!) at about 10^{-2} cd/m^2 (Reichardt, 1973; Wehrhahn, 1976	?
Present in the *Drosophila* mutant S 129 (Heisenberg, pers. comm.)	Disturbed in the *Drosophila* mutant S 129 (Heisenberg, pers. comm.)	?

FIGURE 5

1: "Simple" algorithms

An algorithm operates on some kind of input and yields a corresponding output. In formal terms, an algorithm can be thought of as a mapping between the input and the output space. Perhaps the simplest of all nonlinear operators on a linear space are the so-called polynomial operators. They encompass a broad spectrum of applications including all linear problems, and they approximate all sufficiently smooth, nonlinear operators. For this particular class of "simple" algorithms (i.e. representable through a "smooth" operator) polynomial representations provide a canonical decomposition in a series of simpler, multilinear operators. Figure 5 shows this decomposition in terms of interactions or "graphs" of various orders: in this way an algorithm, or its network implementation, may be decomposed into an additive sequence of simple, canonical terms, just as in another context, a function can be conveniently characterized by its various Fourier terms. Moreover, functional and computational properties can be associated with interactions of a given order and type.

Poggio & Reichardt (1976) used the polynomial representation of functionals to classify the algorithms underlying movement, position and figure-ground computation in the fly's visual system. The idea was to identify which terms, among the diversity of the possible ones, are implied by the experimental data. Figure 5 shows the graphs that play a significant role in the fly's control of flight and, in this sense, characterize the algorithms involved. The notion that seems to capture best the "computational complexity" of these simple, smooth mappings is the notion of p-order (perceptron-order, see Poggio and Reichardt, 1976). Movement computation in the fly is of order 2, and figure-ground discrimination in the simple case of relative motion depends on fourth-order graphs, but possibly with p-order 2. A closed or Type 1 (Marr 1976b) theory of this kind may be a useful way of characterizing preprocessing operations in nervous systems. The approach has a rather limited validity however, since it does not apply to the large and important class of "non-smooth" algorithms, where cooperative effects, decisions and symbols play an essential role. While an arbitrary number of mechanisms and circuits may implement these "smooth" algorithms, it is clear that "forward" interactions between neurons are the most natural candidates.

Although the various levels of description are only loosely related, knowledge of the computation and of the algorithm may sometimes admit inferences at the lowest level of anatomy and physiology. The description of the visual system of the fly at the computational and functional level suggests, for instance, that different, separate neural structures may correspond to the different computations. Recent data support this conjecture. Movement computation (the term $r\dot{\psi}(t)$ of equation (1) and the second order graph of Figure 5) seems to depend mainly on receptor system 1-6, while the position computation (the term $D(\psi(t))$ of equation (1) and the "self-graph" of Figure 5) seems dependent on receptor system 7-8 (Wehrhahn, 1976 and in preparation). Mutants of Drosophila, normal with respect to the movement algorithm, are apparently disturbed in the position algorithm (Heisenberg, in preparation).

2: "Cooperative" algorithms

A more general and not closely definable class of algorithms includes what one might call cooperative algorithms. Such algorithms may describe bifurcations and phase transitions in dynamical systems. An essential feature of a cooperative algorithm is that it operates on many "input" elements and reaches a global organization *via* local but highly interactive constraints. An apparently cooperative algorithm plays a major role in binocular depth perception (Julesz 1971). The stereopsis computation defined by figure 2a applies many local constraints to many local inputs to yield a final state consistent with these constraints. Various mechanisms could implement this type of algorithms. Parallel, recurrent, nonlinear interactions, both excitatory and inhibitory, seem to represent a natural implementation. In the stereopsis case such a mechanism is illustrated in the rest of figure 2. This mechanism may be realized through many different components and circuitries. In the nervous system, however, there are certain very obvious candidates, which allow some definite predictions. For instance, one is led to conjecture the existence of disparity columns (actually layers) of cells with reciprocal excitatory short-range interactions on each layer and long-range inhibitory interactions between layers with the characteristic "orthogonal" geometry of figure 2. Figure 6 shows that this algorithm succesfully extracts depth information from random-dot stereograms. The algorithm exhibits typical cooperative phenomena, like hysteresis and disorder-order transitions. It is important to stress that it is the computational problem which determines the structure of the excitatory and inhibitory interactions, and not "hardware" considerations about neurons or synapses. The apparent success of this cooperative algorithm in tackling the stereo problem suggests that other perceptual computations may be easy to implement in similar ways. Likely candidates are "filling-in" phenomena, subjective contours, figural reinforcement, some kinds of perceptual grouping and associative retrieval. In fact the associative retrieval network described by Marr (1971), in connexion with a theory of the hippocampal cortex, implements a cooperative algorithm.

3: Procedural algorithms

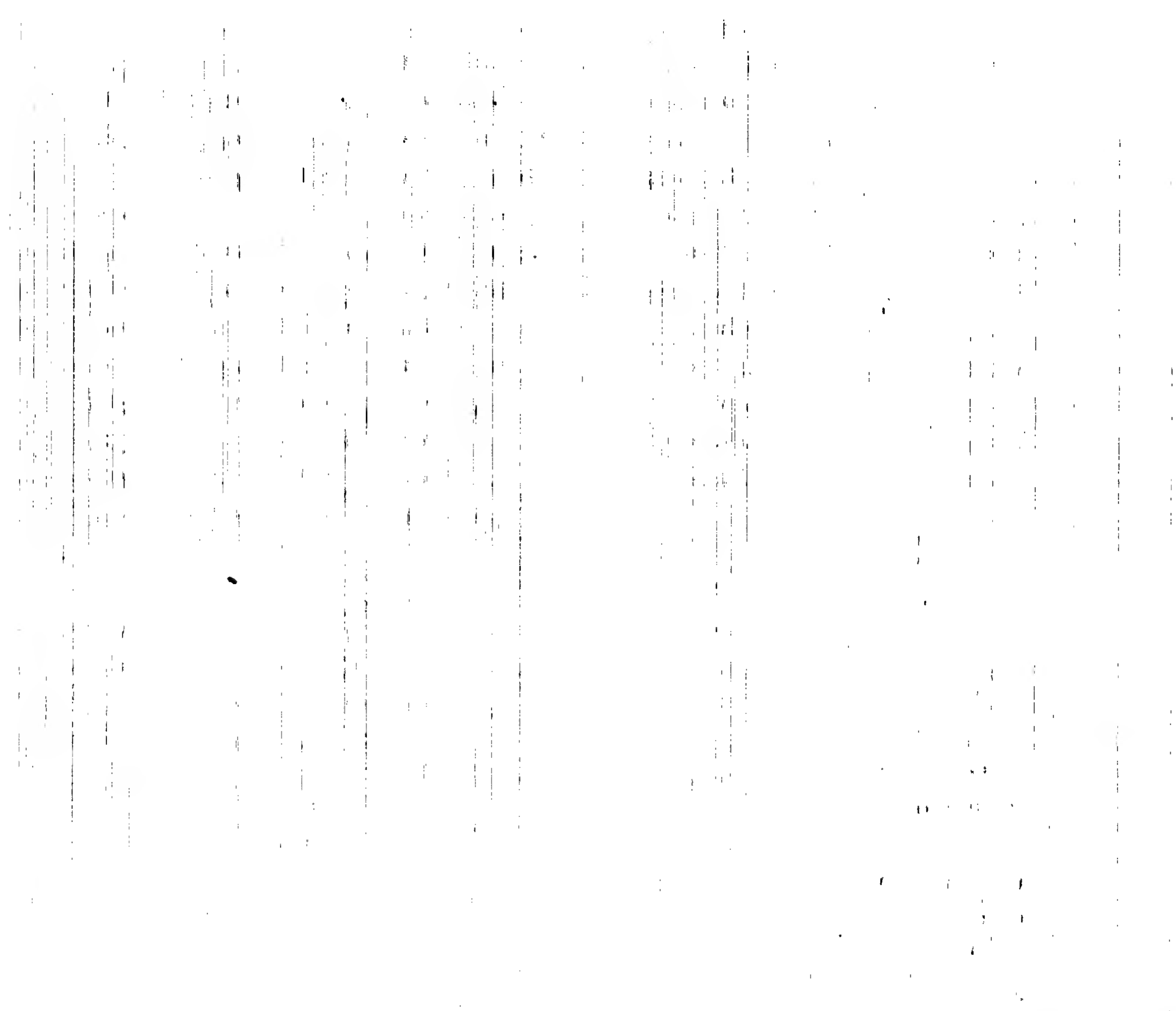
Still another and larger class of algorithms is represented by the specification of procedures, and the construction and manipulation of explicit symbolic descriptions. For example, the 3-D representation theory described in part 5 of the previous section explains how the stick figure representation of a viewed object may be obtained from an image, and manipulated during recognition. The detailed specification of the algorithms involved here is carried out by defining the datastructures that are created to represent the situation, and by specifying procedures that operate on these datastructures in accordance with the information currently being delivered from the image, and that available from stored models.

This way of specifying an algorithm is very general and powerful, although unlike the two other ways that we discussed, it is a far cry from the circuitry level of description at which neurophysiological experiments are carried out. In a digital computer, one does not try to bridge the gap between these two levels in one step. Instead, a basic instruction set, an assembler, a high level language (LISP, ALGOL) and a compiler

Figure 6. A pair of random dot stereograms (left and right), the initial state of a network that implements the algorithm illustrated in figure 2, and the first 9 iterations of the network operating on this stereo pair. To understand how the figures represent states of the network, imagine looking down on it from above. The different disparity layers in the network are in parallel planes spread out horizontally, and the viewer is looking down through them. In each plane, some nodes are on and some are off. Each layer in the network has been assigned a different gray level, so that a node that is switched on in the lowest layer contributes a dark point to the image, and one that is switched on in the top layer contributes a lighter point. Initially (iteration 0) the the network is disorganized, but in the final state order has been achieved (iteration 9). The central square has a convergent disparity of 2 relative to the background, and it therefore appears lighter. The density of the original random dot stereogram was 50%, but the algorithm succeeds in extracting disparity values at densities down to less than 5%. Let C^{n} denote the state of a cell (either 0 or 1) in the 3-D array of fig.2b at the n-th iteration. Then the algorithm used here reads

$$C^{(n+1)}_{ijh} = u_\theta \{ \sum_{S(ijh)} C^{(n)}_{i'j'h'} - \epsilon \sum_{O(ijh)} C^{(n)}_{i'j'h'} + C^{(0)}_{ijh} \},$$

where $u_\theta(x) = 0$ if $x < \theta$, and $= 1$ otherwise; S(ijh) is a neighborhood of cell (ijh) on the same disparity layer; O(ijh) represents the neighborhood of cell (ijh) defined by the "orthogonal" directions shown in Fig. 2b. Excitation between disparity layers is also present in a diagonal direction orthogonal to the layers and decreases with increasing disparity "distances"; D(ijh) represents the corresponding neighborhood.



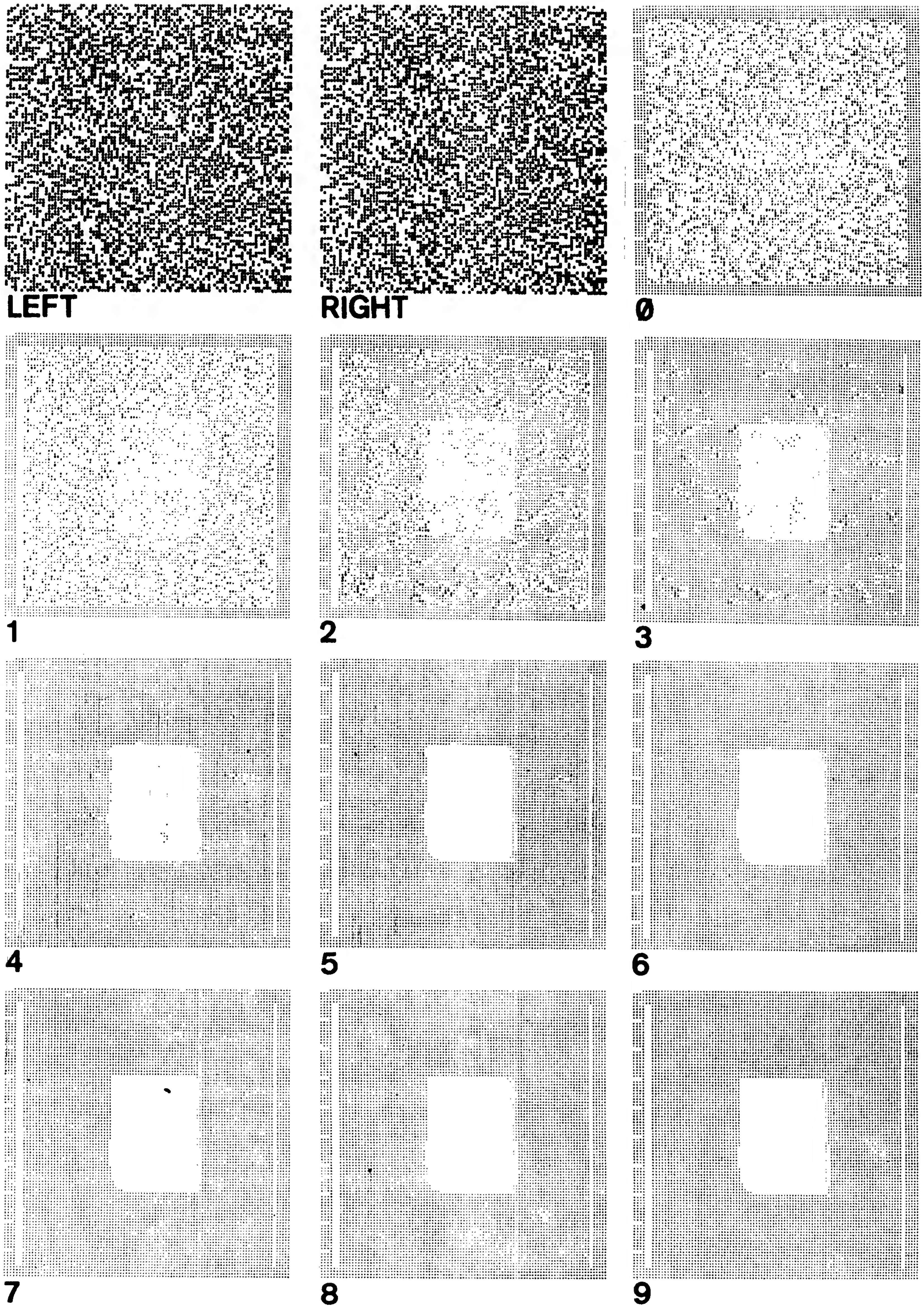


FIGURE 6

are interposed to ease the burden of passing from the description of a computation down to the specification of a particular pattern of current flow.

We may eventually expect a similar intermediate vocabulary to be developed for describing the central nervous system. Hitherto, only one non-trivial "machine-code" operation has been studied in the context of neural hardware, namely simple storage and retrieval functions (Marr 1969 & 1971, and Brindley 1970).

ACKNOWLEDGEMENTS: T. P. acknowledges the financial support of the Max-Planck-Gesellschaft during his visit to M.I.T.

References

Brindley, G. S. (1970). Nerve net models of plausible size that perform many simple learning tasks. *Proc. Roy. Soc. B., 174*, 173-191.

Cooley, J. M. & Tukey, J. W. (1965). An algorithm for the machine computation of complex Fourier series. *Math. Comp., 19*, 297-301.

Geiger, G. & Poggio, T. (1975). The Mueller-Lyer figures and the fly. *Science 190*, 479-480.

Julesz, B. (1971). *Foundations of Cyclopean Perception*. Chicago: The University of Chicago Press.

Marr, D. (1969). A theory of cerebellar cortex. *J. Physiol. (Lond.) 202*, 437-470.

Marr, D. (1971). Simple memory: a theory for archicortex. *Phil. Trans. Roy. Soc. B., 252*, 23-81.

Marr, D. (1974). A note on the computation of binocular disparity in a symbolic, low-level visual processor. *M. I. T. A. I. Lab. Memo 327*.

Marr, D. (1976a). Early processing of visual information. *Phil. Trans. Roy. Soc. B.* (in the press).

Marr, D. (1976b). Artificial Intelligence -- a personal view. *M. I. T. A. I. Lab. Memo 355*.

Marr, D. & Nishihara, H. K. (1975). Spatial disposition of axes in a generalized cylinder representation of objects that do not encompass the viewer. *M. I. T. A. I. Lab. Memo 341*.

Poggio, T. & Reichardt, W. (1976). Visual control of the orientation behavior of the fly: towards the underlying neural interactions. *Quarterly Reviews in Biophysics*, (in the press).

Reichardt, W. (1970). The insect eye as a model for analysis of uptake, transduction and processing of optical data in the nervous system. In: *The Neurosciences, Second Study Program.* (ed. F. O. Schmitt), 494-511. New York: The Rockfeller University Press.

Reichardt, W. & Poggio, T. (1976). Visual control of the orientation behavior of the fly: a quantitative analysis. *Quarterly Reviews in Biophysics*, (in the press).

Warrington, E. K. & Taylor, A. M. (1973). The contribution of the right parietal lobe to object recognition. *Cortex, 9*, 152-164.

Wehrhahn, C. (1976). Evidence for the role of retinal receptors R 7/8 in the orientation behavior of the fly. *Biol. Cybernetics, 21*, 213-220.